

MTrill project: Machine Translation impact on language learning

Natália Resende
ADAPT Centre
School of Computing
Dublin City University

`natalia.resende@adaptcentre.ie` `andy.way@adaptcentre.ie`

Andy Way
ADAPT Centre
School of Computing
Dublin City University

1 Background

Over the last decades, massive research investments have been made in the development of machine translation (MT) systems (Gupta and Dhawan, 2019). This has brought about a paradigm shift in the performance of these language tools, leading to an widespread use of popular MT systems (Gaspari and Hutchins, 2007). Although the first MT engines were used for gist-ing purposes, in recent years, there has been an increasing interest in using MT tools, specially the freely available online MT tools, for language teaching and learning (Clifford et al., 2013). The literature on MT and Computer Assisted Language Learning (CALL) shows that, over the years, MT systems have been facilitating language teaching and also language learning (Niño, 2006). Research also shows the positive role of MT systems in the development of writing skills in English as well as in improving communication skills in English (Garcia and Pena, 2011). However, to date, the cognitive impact of MT on language acquisition and on the syntactic aspects of language processing has not yet been investigated and deserves further scrutiny.

The MTrill project aims at filling this gap in the literature by examining whether MT is contributing to a central aspect of language acquisition: the so-called language binding, i.e., the ability to combine single words properly in a grammatical sentence (Heyselaar et al., 2017; Ferreira and Bock, 2006). The project focus on the initial stages (pre-intermediate and intermediate) of the acquisition of English syntax by Brazilian Portuguese native speakers using MT systems as a support for lan-

guage learning.

Below, we present the methodological approach of the project as well as its objectives.

2 Methodological approach and objectives

In order to examine the impact of MT systems on the acquisition and processing of English syntactic structures, this research will implement a syntactic priming laboratory study to investigate how memory systems encode the syntactic information triggered by MT output. Syntactic priming is an experimental paradigm commonly used by researchers in the field of psycholinguistics as an way to understand aspects of the representation and processing of language syntax (Braningan et al., 2000). The Syntactic priming or structural priming can be defined as the tendency speakers have to use a syntactic structure that has been previously encountered (Bock, 1986).

A common syntactic priming experimental paradigm (Bock, 1986) consists of reading a sentence and asking participants to repeat out loud the same sentence. Listening and repeating the sentence is considered the “prime phase” of the experiment as the experimenter could control participants’ exposure to different syntactic structures. Following the “prime phase”, participants are requested to describe an image so that the researcher could observe if participants would use in the subsequent utterance the same structure they had just produced in the prime phase. In our experiments, we will check whether after being exposed to the output of an MT through a translation task using Google Translate, participants will use the syntactic alternative of the MT output or whether they will choose a different syntactic alternative in their subsequent speech. The language pair investigated

is Portuguese–English due to the differences in syntactic structures between them.

The specific objectives of the project are:

- Investigate whether MT systems are capable of eliciting syntactic priming effects.
- Investigate whether any priming effect elicited is of explicit or implicit learning nature.
- Investigate whether the size of the priming effect varies as a function of the gender of the subjects.
- Investigate whether the size of the effect varies as a function of the level of reliability of system or English proficiency levels.

3 First study

In the first study of the MTrill project, 20 Brazilian Portuguese speakers at intermediate and advanced English proficiency levels took part in syntactic priming experiment that investigated the influence of the popular Google Translate MT system on the mental processing of English as a second language.

We adopted a pretest-priming design (Shin and Christianson, 2012). In the pretest phase, participants were instructed to freely translate from Portuguese into English sentences depicting images appearing in a computer screen. Immediately after this pretest, in the priming phase, participants were instructed to translate other sentences depicting images appearing in a computer screen using Google Translate and repeat out loud the output. After this task, they were requested to describe out loud an image using three words provided. This design allowed to test if participants would describe the image mirroring the syntactic structure previously seen in the output of Google Translate and compare participants' language behaviour before and after Google Translate.

The results of this preliminary study show that, after performing a translation task with Google Translate, participants described images in English using more frequently the syntactic alternative previously seen in the output of Google Translate when compared to the translation task with no prior influence of the MT output. Results also show that this priming effect is modulated by language proficiency levels.

Acknowledgements

The ADAPT Centre for Digital Content Technology (www.adaptcentre.ie) at Dublin City University is funded by the Science Foundation Ireland Research Centres Programme (Grant 13/RC/2106) and is co-funded by the European Regional Development Fund. This project was funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement number 843455. Start date 25/04/2019 - End date 24/04/2021.

References

- Bock, J.Kathryn. 1986. Syntactic persistence in language production. *Cognitive Psychology*, 18(3):355 – 387.
- Braningan, Holly P., Martin J. Pickering, Andrew J. Stewart, and Janet F. Mclean. 2000. Syntactic priming in spoken production: Linguistic and temporal interference. *Memory and Cognition*, 28(8):1297–1302.
- Clifford, Joan, Lisa Merschel, and Joan Munné. 2013. Surveying the landscape: What is the role of machine translation in language learning? *Revista d'innovacio educativa*, (10):108–121.
- Ferreira, Victor S. and Kathryn Bock. 2006. The functions of structural priming. *Language and Cognitive Processes*, 21(7-8):1011–1029. PMID: 17710210.
- Garcia, Ignacio and María Isabel Pena. 2011. Machine translation-assisted language learning: writing for beginners. *Computer Assisted Language Learning*, 24(5):471–487.
- Gaspari, Federico and John Hutchins. 2007. Online and free ! ten years of online machine translation : Origins , developments , current use and future prospects.
- Gupta, Brij M. and S. M. Dhawan. 2019. Machine translation research a scientometric assessment of global publications output during 2007 16. *DESI-DOC Journal of Library Information Technology*, 39(1):31–38, 01.
- Heyselaar, Evelien, Peter Hagoort, and Katrien Segaert. 2017. In dialogue with an avatar, language behavior is identical to dialogue with a human partner. *Behavior Research Methods*, (49):46–60.
- Niño, Ana. 2006. *Evaluating the suitability of using raw machine translation output as input for foreign language written production*. Ph.D. thesis.
- Shin, Jeong-Ah and Kiel Christianson. 2012. Structural priming and second language learning. *Language Learning*, 62(3):931–964.